Quantile Regressions and Quantile Treatment Effects

- QTE allows us to study the heterogeneity in treatment effects.
- Possible for ATE = 0 and QTE < 0 or QTE > 0

$$QTE(\tau) = F_1^{-1}(\tau) - F_0^{-1}(\tau) = Q_1(\tau) - Q_0(\tau)$$

- QTE can be estimated using quantile regressions. Quantile regressions estimate the conditional quantile function (analogous to regression estimating the ATE).
- The conditional quantile function Q(Y|X) is analogous to the conditional expectation function which is E(Y|X). With OLS regression we parametrize E(Y|X), quantile regression instead parametrizes the conditional quantile function Q(Y|X).

$$\hookrightarrow Q_\tau(Y_i|X_i) = F_Y^{-1}(\tau|X_i)$$

$$\hookrightarrow \tau = 0.5 \Rightarrow Q_\tau(Y_i|X_i) = med(Y_i|X_i)$$

$$\hookrightarrow Q_\tau(Y_i|X_i) = X_i\beta_\tau \Rightarrow Y_i = X_i\beta_\tau + \varepsilon_i, \quad Q_\tau(\varepsilon_i|X_i) = 0$$

$$\hookrightarrow \hat{\beta}_\tau = \underset{\beta}{\arg\min} \frac{1}{n}\sum_{i=1}^{n} \rho_\tau(y_i - X_i\beta), \quad \rho_\tau(u) = (\tau - I(u \le 0))u$$

$$\hookrightarrow \text{Assym. loss} = \begin{cases} \tau(y_i - X_i\beta) & \text{if } y_i > X_i\beta \\ (\tau-1)(y_i - X_i\beta) & \text{if } y_i \le X_i\beta \end{cases}$$

$$\hookrightarrow \tau = 0.5 \Rightarrow \hat{\beta}_{0.5} = \underset{\beta}{\arg\min} \frac{1}{n}\sum_{i=1}^{n} |y_i - X_i\beta|$$

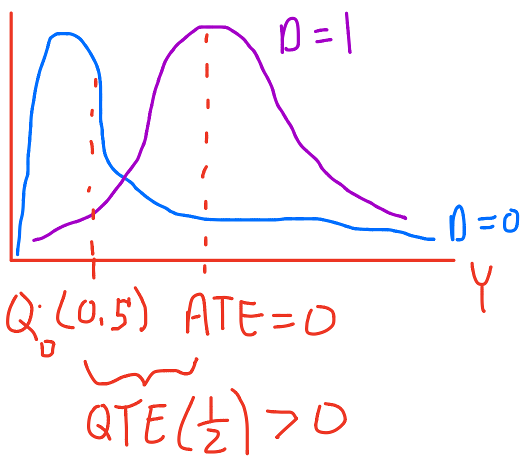
- Hence QTE can estimate using the quantile regression as shown below:

$$Q(\tau|X_i) = \alpha(\tau) + \beta(\tau)X_i$$

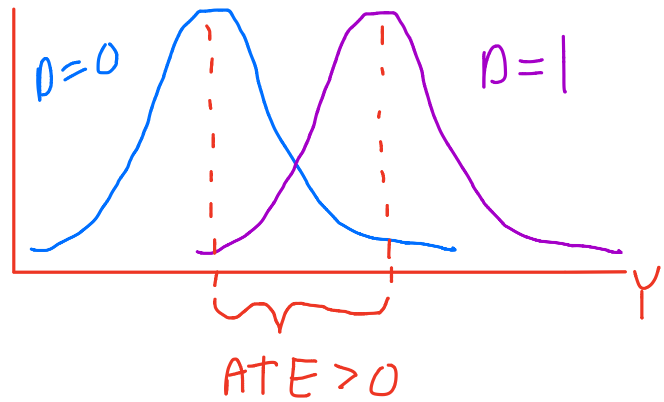
$$\hookrightarrow \begin{cases} \alpha(\tau) = Q_0(\tau) \\ \beta(\tau) = Q_1(\tau) - Q_0(\tau) \end{cases}$$

- Constant coefficient estimates across quantiles if there is just a pure location shift due to the treatment. QTE = ATE in case of just location shift
- Distribution of outcome may change in ways that are not explained by mean changes. This is where quantile regression and QTE have added value over standard regression.

Distributional Change



Location Shift



- If intervention is rank preserving, and we have positive coefficients at the lower decile, than those with lower outcome now have higher outcomes under intervention.
- Conditional QTE are defined in terms of the regression values, whereas unconditional QTE is the effect of the treatment on the entire population.
- Conditional QTE assume the treatment is exogenous conditional on X. Unconditional QTE is useful when treatment is exogenous.
- The seminal Keener and Basset (1978) paper is for conditional QTE. If no covariates are includes, the conditional QTE coincide with unconditional QTE.

<u>POM</u>: $Y_i = Y_{i1} D_i + Y_{i0}(1 - D_i)$

$$\hookrightarrow \begin{cases} Y_{i1} = X_i \beta_\tau + D_i \delta_\tau + \varepsilon_i \\ Y_{i0} = X_i \beta_\tau + \varepsilon_i \end{cases}, \quad Q_\tau(\varepsilon_i \mid X_i) = 0$$

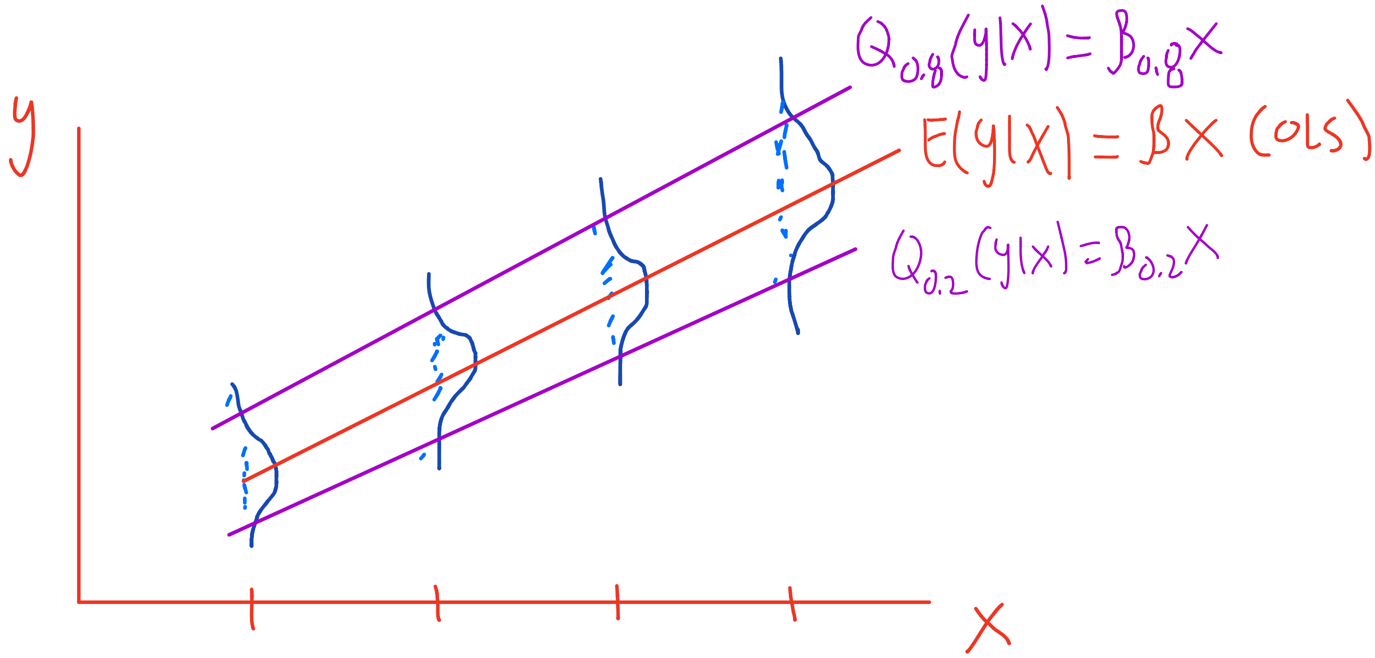
$$\hookrightarrow Q_\tau(Y_i \mid X_i, D_i) = X_i \beta_\tau + D_i \delta_\tau$$

$$\hookrightarrow \text{Cond. QTE} = \delta_\tau$$

- Conditional QTE are identified from the joint distribution of (Y, X, D)

$$(\hat{\beta}_\tau, \hat{\delta}_\tau) = \underset{(\beta, \delta)}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^{n} \rho_\tau (Y_i - X_i \beta_\tau - D_i \delta_\tau)$$

- Quantile regression is not the same as doing least squares estimation on segments of the outcome variable. The objective function in quantile regression relies on the check function.
- The Figure below compares OLS and quantile regression:



- Note that OLS goes fits the mean of the conditional distributions f(y|x). Whereas the quantile regression fits the corresponding quantile of f(y|x). Since in the above figure the conditional distributions have the same mean and median, the 0.5 quantile regression will align with OLS regression.